

about corrections they believe are necessary, and provide feedback to those officials to document our actions on each suggested correction. This includes an opportunity to review the updated maps and address listings after processing the corrections.

**Maintaining the quality of our Geographic Support products:** we employ statistically sound techniques to the quality evaluations that build, update, or enhance the TIGER database and the Master Address File. Quality evaluation and correction is an integral part of our operations, ensuring that any sub-standard work is identified at the earliest possible stages.

**Acquiring TIGER and Master Address File hardware:** acquiring replacement hardware to process the TIGER database is critical to the success of all Geographic Support efforts. The hardware will help produce maps needed to implement our cooperative programs, update the TIGER and Master Address File databases, and provide acceptable products for new partners in our cooperative programs.

The Geography program area supports Decennial Census activities and provides accurate geographic information and lists of spatially located addresses. These are essential for the U.S. Census Bureau to provide accurate statistical data for:

- apportioning Congressional seats among the states;
- redistricting Congressional and state legislative representation and governmental or administrative subdivision;
- distributing federal and state funds for formula grant programs; and
- analyzing economic and demographic data by private, academic, and government sectors.

The Geography Division plans, coordinates, and administers all geographic and cartographic activities needed for the U.S. Census Bureau's nationwide statistical programs; Geography also manages the U.S. Census Bureau's continuous updates of features, boundaries and geographic entities.

The Geography Division works with state, local, and tribal governments in helping to maintain national spatial data. The Geography Division supports the efforts of the Federal Geographic Data Committee under Executive Order 12906. The U.S. Census Bureau uses addresses and related information from the U. S. Postal Service in exchange for geographic and spatial information (Public Law 103-430).

## 1.2 Geography Division IT Objectives

The Geography program area's IT objectives are:

- providing geographic data and supporting customers inside and outside of the U.S. Census Bureau;
- exploring new technologies in digital cartography and Geographic Information Systems;
- reducing software development costs by using commercial off-the-shelf software when possible; and
- continuing our cooperative programs with the other federal agencies under the Federal Geographic Data Committee, and with state, local, and tribal governments.

## 2.0 Geography Division IT Support

The table below presents our major geographic IT support systems. Several of these, such as the TIGER database and the Master Address File, support the Decennial Census and the American Community Survey; direct relations to Decennial Census activities are indicated in the second column. The Geography Division also supports (in terms of continuous operations) the Regional Offices and other censuses (e.g., Economic); these continuing operations

are indicated in the third column. A detailed description of each follows the table. In addition to the direct Geographic Support System (GSS) program, GSS also helps provide specialized products and services for the Decennial Census. Please see the Decennial program area section of this Operational IT Plan for the documentation of the development and production of these decennial-specific products and services.

Geography Program Area IT Support	Decennial Related	Continuing Operations
<b>Geographic Support Programs</b>		
Geographic Update System (GUS)	✓	✓
Map Production System	✓	✓
Intranet Geography Web System	✓	✓
Enterprise Support	✓	✓
TIGER	✓	✓
Research Programs	✓	✓
Master Address File (MAF)	✓	✓
Local Update of Census Addresses (LUCA)	✓	
Boundary and Annexation Survey (BAS)	✓	✓
<b>Decennial Support Programs</b>		
Census 2000 Redistricting Data Program	✓	
Participant Statistical Areas Programs	✓	
Urbanized Area Delineation	✓	

## 2.1 Geography Division IT System Descriptions

### 2.1.1 Geographic Support System

To support its various censuses and sample surveys, the U.S. Census Bureau has developed an integrated and automated computer-based Geographic Support System (GSS). The GSS requires large volumes of information from many external sources to establish and maintain accurate geographic boundaries within the TIGER database and accurate addresses in the Master Address File. This information must be updated on a regular basis to validate all of the boundaries and addresses contained in the GSS. GSS provides a geographic framework from which the U.S. Census Bureau can generate the maps for:

- collecting and validating boundaries (government and physical);
- compiling and validating address lists;
- collecting field data; and
- disseminating data products.

GSS also provides a geographic basis for:

- assigning the housing units, farms, and business establishments in each census and sample survey to a specific geographic location;
- automating questionnaire check-in and control systems to generate follow-up assignments or reminder notices; and
- developing a source file from which the U.S. Census Bureau can generate geographic table stubs and summary cartographic products for data tabulation and customer support purposes.

Subsystems within the GSS are:

#### Geographic Update System

The Geographic Update System can reference a portion (such as a single county) of the TIGER database and the Master Address File for updates and corrections to these databases, or for research purposes. It is also used in various areas of the U.S. Census Bureau for reviewing counts, resolving problems, and analyzing data.

#### Map Production System

Map production is done at the Regional Census Centers, the National Processing Center, and Headquarters. The maps (in digital form as a Map Image Metafile) are then transmitted to the actual map plotting sites. Software then converts the Map Image Metafiles to output device-specific format (Hewlett-Packard Graphical Language or PostScript) and sends the resulting print file to the output device. Thus, map production is a two-step process. First, on a computer system where a Master Address File and TIGER database is located, a Map Image Metafile is created (the Map Image Metafile may contain more than one map sheet). The Map Image Metafile may then be processed for printing immediately, copied by network to another computer system for printing later, or copied to CD-Recordable for shipment to a remote printing site (for example, Local Census Offices). The following are typical map production methods:

## Topologically Integrated Geographic Encoding and Referencing (TIGER ) System

TIGER supports all Demographic and Economic current surveys, the Economic Censuses, various research and development projects, the Decennial Census pretests, various Census 2000 operations, and the American Community Survey.

The TIGER database is divided into 3,287 county partition files, with some counties further subdivided into multiple partitions to improve processing efficiency. Updates and corrections to the TIGER database and Master Address File take place in the Regional Census Centers, the National Processing Center, and at Headquarters, using graphic workstations and PCs connected to file servers at their location. Database partitions are checked out from the master database at the Bowie Computer Center for update operations.

Corrections and additions to the TIGER database are made by clerks using annotated paper copies of census maps or by computer-assisted transfer of new feature information from locally provided spatial files.

As part of keeping TIGER up-to-date, the U.S. Census Bureau has partnerships with local governments. Local governments and regional and metropolitan agencies assist us in locating and updating missing or incorrect street

features, street names, and address ranges. This information is used to link the Master Address File to the TIGER database.

The TIGER Control System controls the movement of the TIGER and Master Address File files between processing sites (Regional Census Centers, Headquarters, the National Processing Center, etc.), tracks their locations, tracks processing activities, and ensures database integrity. The decentralized approach to updating and validating the TIGER database (by providing the relevant file locally to the Regional Census Centers for processing) must be coordinated with other outside updates and validations so that database integrity is preserved. This system uses considerable computer resources for integrity edits as well as significant telecommunications bandwidth.

Figure 1, shown on the following page, illustrates the Geographic Support System update process for partnership programs.

### **Research Programs**

TIGER has several shortcomings; one of these is its lack of positional accuracy. The U.S. Census Bureau has looked into various projects to improve the coordinate quality of the TIGER database: two of these projects currently being investigated are the use of the Global Positioning System and the use of digital imagery by the Geography Division's Geospatial Research and Standards Staff.

The Global Positioning System is a satellite-based navigation and real-time transfer system; by using its receiver, a user is able to identify their exact location and elevation anywhere on Earth. Because this system is highly accurate and reliable system, the Geospatial Research and Standards Staff is considering using this technology to improve the TIGER database. This would be achieved by using the Global Positioning System to accurately locate various roads, housing units, etc. that exist both in the real world and in TIGER. Similarly, the coordinates of roads and other features will be captured from digital images of the same areas. Once these data are gathered, they would be used to correct errors in TIGER. Ultimately, this could produce a TIGER database that will not only be current, but will also be highly accurate.

### **Master Address File**

The Master Address File is a database of all addresses in the United States; the residential addresses in this database will be validated prior to Census 2000 mailing and enumeration activities. It is

an essential component of every census operation in which housing unit addresses are used. For example, addresses in the Master Address File are critical to the production of maps. Thus, the Master Address File can be thought of an address database with geographic components that complements the TIGER database, a geographic database with address components. Programs and projects that rely on detailed address information use the Master Address File.

### **Local Update of Census Addresses (LUCA)**

LUCA offers local governments the opportunity to review and update the list of individual residential addresses used for Census 2000. This list is based on the addresses in the Master Address File. Public Law 103-430 calls for this review to be conducted by locally designated census liaisons who must agree to maintain the confidentiality of the information on the census address list. The U.S. Census Bureau Regional Offices will maintain contact with the liaisons and provide feedback on the making of submitted corrections in accordance with Public Law 103-430.

### **Boundary and Annexation Survey**

The U.S. Census Bureau normally conducts a Boundary and Annexation Survey (BAS) each year to maintain the inventory of general purpose governments and to update information about the boundaries of larger population entities. In years ending in 1,3,4,5, and 6, we limit the BAS to incorporated places with a population of 5000+. In years ending in 2 and 7, we include in

Components of the Geographic Support System are located in the following areas.

**Bowie Computer Center:** the master TIGER database and the Master Address File reside on a single Open VMS Alpha cluster currently consisting of two systems.

**Regional Offices/Regional Census Centers:** users update and access the TIGER and Master Address Files databases and are able to print maps and address lists. At the time the Regional Census Centers were opened, the geographic IT equipment was transferred from the Regional Offices to the Regional Census Centers. The equipment will be returned after the Regional Census Centers are closed. Each Regional Census Center/Regional Office has UNIX file servers, graphic workstations, PCs, and printers.

**Suitland Beta Site:** the replica of the complete Regional Census Center configuration includes geographic operations for testing the configuration software distribution to the Regional Census Centers.

**National Processing Center:** same as our Regional Offices. The emphasis is on continuing census operations.

**Headquarters:** Headquarters performs all software development work; major system administration for itself, the Regional Census Centers, and the National Processing Center; and coordination with other federal agencies. Headquarters also has users that update and access the TIGER and Master Address File databases and are able to print maps and address lists. The

emphasis is on special purpose and management operations. Headquarters also has a replica of the Regional Census Center geographic configuration for preliminary software testing.

Dedicated data communication links are used for file transfers and sending control/status information between the Bowie Computer Center, the Geography Division, the National Processing Center, the Regional Census Centers, and the Regional Offices. Local Area Networks connect graphic workstations and PCs at each site.

U.S. Census Bureau employees in all the Regional Offices update and correct the TIGER database and the Master Address File by using graphic workstations and PCs connected to UNIX file servers. Those UNIX file servers are linked via the U.S. Census Bureau's Wide Area Network to the main file server/processing environment in the Bowie Computer Center.

Corrections and additions to the TIGER database are made by clerks using annotated paper copies of census maps or by computer-assisted transfer of new feature information from locally provided spatial files.

The map production system creates maps at all Regional Census Centers, Regional Offices, and the National Processing Center. Map Image Metafiles, a digital construct of a map, are produced on file servers and then converted to Hewlett-Packard Graphical Language or Adobe PostScript, then transmitted, depending on map type, to small- or large- format printers.



### Geography Distributive Production Clusters

We have found that a cost-effective means of getting work done in the Regional Census Centers is to use multiple application servers that are accessed by PCs using X-windows emulation software. We have also found that this configuration is more cost-effective than graphics workstations/disk farms or a combination of graphics workstations/disk farms and PCs. With the shift to PCs/application servers, we have reduced the original hardware cost estimate for supporting the Regional Census Centers.

- Regional Census Center/Beta Site/National Processing Center

Included are multiple Silicon Graphics UNIX multi-processor file servers with at least 100 GB disk storage for map and TIGER/MAF update production work. Associated with these servers are variable numbers (based on production requirements) of PCs and Silicon Graphics graphic workstations.

- Headquarters
  - Multiple Silicon Graphics UNIX multi-processor file servers with at least 200 GB disk storage for automated processing of local/commercial digital spatial data files, support for Oracle applications, backup services, and Intranet applications;
  - Multiple Windows NT, Novell, and UNIX servers to support office automation ap-

plications such as e-mail, scheduling, Primavera, SAS, the Geographic Information System, and Oracle. Associated with these servers are Intel and Motorola based PCs; and

- a Silicon Graphics UNIX multi-processor file server for Data Access and Dissemination System support.

### Research Server

- Silicon Graphics UNIX multi-processor file server with at least 100 GB disk storage and 1 GB memory;
- Software development for the UNIX machines and UNIX workstations in the Regional Census Centers, the National Processing Center and at Headquarters, using ANSI C, C++, Tcl/Tk, and Perl. All relational applications will use the U.S. Census Bureau standard Oracle database.

Other devices include:

- **Map output devices:** support boundary collection and validation activities, to build and maintain the linked TIGER / Master Address File databases, and to let state and local officials preview the status of information in the TIGER database by providing them with maps they can update and return to the U.S. Census Bureau.

## 2.1.2 Geographic Support System Progress Against Planned Milestones

Geographic Support System Milestones, FY 98					
Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Demographic Census Entities: Military Installations	04/97	03/98		03/98	Completed.

Geographic Support System Milestones, FY 99					
Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Geographic Services – Participant Statistical Entities: 1998 Boundary and annexation survey	06/97	10/98		10/98	Completed.
Redistricting Program: Phase I-Block Boundary Suggestion Program	04/94	03/99		03/99	Completed.
Upgrade and replace database and application servers for 1999	12/98	03/99		03/99	Completed.
21 <sup>st</sup> Century MAF/TIGER Initiative Business Case	01/99	05/99		05/99	Completed.
Local Update of Census Address (LUCA): Design appeals Process-Decennial LUCA	10/96	06/99		06/99	Completed.
Create MAF for Non-City Style Address Areas: Implement Address Listing	10/96	07/99			In progress.
MAF Improvement for City-Style Address Areas: Block Canvassing	09/97	07/99			In progress.
Obtain desktop graphic workstations for research work for 1999	07/99	08/99			
Obtain large format map printer (E-size)	08/98	09/99			In progress.
Obtain additional laser and inkjet printers to support TIGER/MAF production efforts	03/99	09/99			In progress.



### Geographic Support System Milestones, FY 01

Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Redistricting Program: Phase 2-Voting Districts/Legislative Districts	04/97	10/00			In progress.
TIGER/MAF/DMAF Linkages/Transactions	11/96	11/00			In progress.
Special Geographic Products	01/96	01/01			In progress.
Geographic Comparability Files	11/99	01/01			
Determine 2002 upgrade and replacement requirements for database and application servers	10/00	01/01			
Upgrade and replace database and application servers for 2002	10/00	01/01			
TIGER/Line Products	08/98	02/01			In progress.
Tract/Voting District Maps	08/98	03/01			In progress.
County Subdivision Maps	08/98	03/01			In progress.
County Block Maps	08/98	07/01			In progress.
Reference Maps for 2000 Tabulations Geography	09/98	07/01			In progress.
Boundary Files for 2000 Tabulation	01/99	07/01			In progress.
Statistical Maps for 2000 Tabulation Geography	05/99	07/01			In progress.

### Geographic Support System Milestones, FY 02

Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Demographic census Entities: Urban/Rural	04/97	11/01			In progress.
Determine 2003 upgrade and replacement requirements for database and application servers	10/01	01/02			
Upgrade and replace database and application servers for 2003	10/01	01/02			
Boundary Files for Post – 2000 Demographic Entities	10/01	02/02			
Reference Maps for Post – 2000 Demographic Entities	10/01	09/02			

The Geographic Products Quality Assurance Team designs and implements a quality assurance program that ensures that all Geography Division's products meet the sponsor's specifications. As necessary, the team assists the Division's branches and staffs in the quality assurance process; in addition, the team also does the following:

- develops quality assurance procedures that are implemented on internal processes;
- works with Geography offices to design methods to assess and report the quality of TIGER and Master Address File products and data; and
- works with the Decennial Statistical Studies Division to ensure that the quality assurance processes are statistically-based and conform to U.S. Census Bureau standards;

The team will design a reporting system that captures the results of all implemented quality assurance activity.

#### 2.1.4 Geographic Support System Risks

We foresee staffing as one of our major risks. Hiring and retaining competent technical staff is becoming increasingly difficult in the federal government. Our recent experience in attempting to obtain qualified computer programmers indicates that the federal government salary levels are significantly below that of the private sector. To address our inability to attract qualified computer programmers, we are making more extensive use of contract programmers. With more frequent turnover of contract staff, use of contract programmers introduces additional costs to continue to train them on the unique features of the TIGER /Master Address File system. Use of task-based contract programmers may also increase the overall cost of developing and maintaining the software system. We have accounted for the increased use of task-based contract programmers in our current budget estimates.

We are working on business continuity plans to address a number of risks that, if they occurred, could conceivably disable the Geographic Support System. In 1998, the IT Directorate led the effort to create a continuity plan for the Bowie Computer Center, where the Geographic Support System OpenVMS computer cluster is located (along with other large-scale systems used by other U.S. Census Bureau program areas). Based on this example, we are developing continuity plans by site location rather than by system, such as the Geographic Support System. We identified the following site locations as needing continuity plans:

- the Regional Census Centers;
- the National Processing Center; and
- Headquarters.

### 3.0 Geography Division Infrastructure

Desktop computers, although used for office automation, are an integral resource of the Geographic Support System. The desktop computers are used extensively for Geographic Support System production, usually as an X-window server connected by the network to a UNIX server running applications such as GUS-X. Other times, the desktop computers run commercial off-the-shelf Geographic Information Systems

software (ArcInfo, ArcView, Mapitude, or MapInfo) to analyze the Master Address File, TIGER, or related information.

The Geography Division manages and supports its own desktop computers and network operating systems through the Computer Support Branch in the Geography Division.

#### 3.1 Geographic Support System Infrastructure Description

The Geography Division's Geographic Support System infrastructure consists of:

- one Novell Netware 4.11 server (Compaq at 200 MHz);
- four Microsoft Windows NT 4.0 servers (one Dell at dual 200 MHz, two Dells at 400 MHz, one Midwest Micro at 233 MHz);
- one Appleshare 4.1 server (Apple at 150 MHz);
- one SMTP e-mail server (one Silicon Graphics at 225 MHz);
- two calendaring servers (one Apple at 150 MHz, one Hewlett Packard);
- 282 Microsoft Windows 95 and NT desktop computers;
- 40 Apple Macintosh 8.1 and 8.5 desktop computers;
- 31 laser printers;
- four inkjet printers; and
- one pin printer.

In addition, approximately 60 desktop computers are used for software development for the Geographic Support System.

We are constantly improving the GSS infrastructure. Within the year, we expect the infrastructure to consist of:

- one Novell Netware 4.11 server (Compaq at 200 MHz);
- six Microsoft Windows NT 4.0 servers (six Dells at dual 500 MHz);
- one SMTP e-mail server (Silicon Graphics at dual 300 MHz);
- one calendaring server (SUN at 175 MHz);
- 290 Microsoft Windows 95 and NT desk-top computers;
- 40 Apple Macintosh 8.5 desktop computers;
- 35 laser printers; and
- six inkjet printers.

Ethernet connects all of the Geographic Support System infrastructure network devices to the U.S. Census Bureau's Wide Area Network. Five virtual Local Area Networks were installed in August 1996 when the Geography Division moved to Washington Plaza (WP1); since then, the IT Directorate has upgraded the router. This year, the IT Directorate is upgrading our connection

**Geographic Support System Infrastructure Milestones, FY 01**

Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Determine PC replacement requirements for 2002	10/00	01/01			
Obtain PC replacement requirements for 2001	01/01	09/01			

**Geographic Support System Infrastructure Milestones, FY 02**

Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Determine PC replacement requirements for 2003	10/01	01/02			
Obtain PC replacement requirements for 2002	01/02	09/02			

**Geographic Support System Infrastructure Milestones, FY 03**

Description	Estimated		Actual		Progress to Date
	Start Date	Finish Date	Start Date	Finish Date	
Determine PC replacement requirements for 2004	10/02	01/03			
Obtain PC replacement requirements for 2003	01/03	09/03			

**3.3 Geographic Support System Infrastructure Performance Measures**

A typical performance measure for infrastructure staff is help desk responsiveness. It was determined during the writing of last year's Operational IT Plan that we should purchase and implement more rigorous help desk software that would track and analyze the responsiveness of the Geography Division infrastructure staff. This project is in the investigative stage; implementation is still expected to be this year.

We have added a performance measure to the Geographic Support System Infrastructure to respond to help desk requests\*. We will begin reporting once we have implemented the help desk.

\*Since our help desk supports both production and office automation IT resources, this performance goal is also included in section 2.3.1 Geographic Support System Performance Measures.